Novel colon-available triterpenoids identified in raspberry fruits exhibit antigenotoxic activities in vitro

McDougall, Gordon J.; Allwood, J. William; Pereira-Caro, Gema; Brown, Emma M.; Verrall, Susan; Stewart, Derek

Published in:
Molecular Nutrition & Food Research

DOI:
10.1002/mnfr.201600327

Publication date:
2017

Document Version
Peer reviewed version

Novel colon-available triterpenoids identified in raspberry fruits exhibit anti-genotoxic activities in vitro.

Gordon J. McDougall[*1], *J. William Allwood*[1], *Gema Pereira-Caro*[2], *Emma M. Brown*[3], *Susan Verrall*[1], *Derek Stewart*[1,4], *Cheryl Latimer*[3], *Geoff McMullan*[3], *Roger Lawther*[5], *Gloria O'Connor*[5], *Ian Rowland*[6], *Alan Crozier*[7] *and Chris I. R. Gill*[3]

[1]Environmental and Biochemical Sciences Group, Enhancing Crop Productivity and Utilisation Theme, The James Hutton Institute, Invergowrie, Dundee DD2 5DA, Scotland

[2]Postharvest, Technology and Agrifood Industry Area, IFAPA, Córdoba, Spain

[3]Northern Ireland Centre for Food and Health, Centre for Molecular Biosciences, University of Ulster, Cromore Road, Coleraine, N. Ireland, UK

[4] NIBIO, Norsk Institut for Bioøkonomi , Pb 115, NO-1431 Ås,

[5]Altnagelvin Area Hospital, Western Health and Social Care Trust, Glenshane Road, Londonderry

[6]Hugh Sinclair Unit of Human Nutrition, Department of Food and Nutritional Sciences, University of Reading, P.O. Box 226, Whiteknights, Reading, UK

[7]Department of Nutrition at the University of California, Davis, California 95616, USA

Correspondence: Gordon J. McDougall

Email: Gordon.mcdougall@hutton.ac.uk

FAX: +44 (0)844 928 5429; **TEL:** +44 (0)1382 568782

Received: 14/04/2016; Revised: 23/08/2016; Accepted:25/08/2016

This is the peer reviewed version of the following article: 'Novel colon-available triterpenoids identified in raspberry fruits exhibit anti-genotoxic activities in vitro', Molecular Nutition & Food Research, which has been published in final form at http://dx.doi.org/10.1002/mnfr.201600327. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Abbreviations: CDDO = 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid; FAV = fruit and vegetables; GI = gastrointestinal; TRF = triterpenoid-rich fraction;

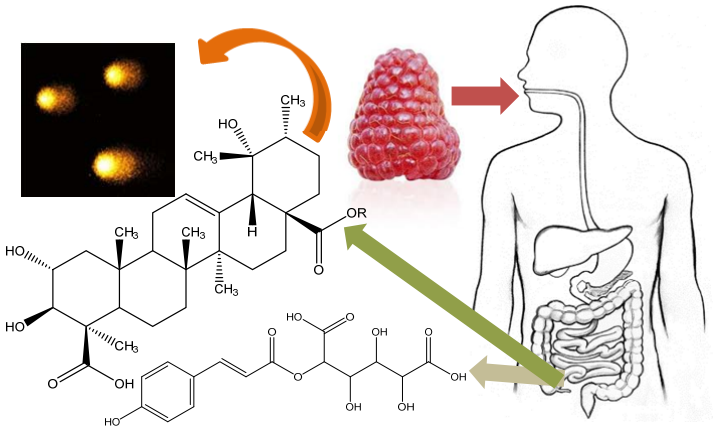
Key words: Raspberries, metabolic profiling, triterpenoids, colon cancer, ileal fluid

ABSTRACT

Scope: Ileostomy studies provide a unique insight into digestion of food, allowing identification of physiologically relevant dietary phytochemicals and their metabolites important to gut health. We previously reported the consistent increase of components in ileal fluids of ileostomates after consumption of raspberries with use of non-targeted LC-MS^n techniques and data deconvolution software highlighting two major unknown components (*m/z* 355 and 679).

Methods and results: In-depth LC-MS^n analyses suggested that the ileal *m/z* 355 components were *p*-coumaroyl glucarates. These compounds have not been identified previously and were confirmed in raspberry extracts after partial purification. The major ileal component with *m/z* 679 was a glycoside with an aglycone of *m/z* 517 and was present as two peaks in extracts of whole puree, unseeded puree and isolated seeds. These components were purified using Sephadex LH20 and C18 SPE units and identified as major, novel raspberry triterpenoid glycosides. This triterpenoid-enriched fraction (100 nM) protected against H_2O_2-induced DNA damage in both colon cancer and normal cell lines and altered expression of cytoprotective genes.

Conclusion: The presence of these novel raspberry triterpenoid components in ileal fluids indicates that they would be colon-available in vivo, so confirmation of their anti-cancer bioactivities is of key physiological relevance.



Two major unknown components present in ileal fluids after intake of raspberries were identified as triterpenoid glycosides and *p*-coumaroyl glucarates using LC-MS^n analysis. The triterpenoids were purified from raspberries and were shown to have genoprotective effects at sub-micromolar levels against H_2O_2–induced damage in colon cell lines. Therefore these novel triterpenoids combine established colonic bioavailability with potential bioactivity at physiological doses.

1 INTRODUCTION

There is great interest in the possibility that non-nutritive components could directly contribute to the health benefits attributable to a diet rich in fruit and vegetables (FAV) [1-3]. One set of phytochemicals of specific interest are the (poly)phenols as are found in high concentrations in certain beverages and FAV, such as berries, and have been attributed a range of important bioactivities [4-9]. However, an important issue with ascribing any potential health benefit is the long held view that the bioavailability of many (poly)phenol classes is low. For example, urinary recoveries of anthocyanins, which reflect passage through the circulatory system, are typically <<1% of intake [10], although recent studies with [^{13}C]cyanidin-3-*O*-glucoside [11, 12] and raspberries [13], which took into account metabolites and the parent anthocyanins, established that cyanidin-based anthocyanins are much more bioavailable than previously envisaged. Likewise with orange juice flavanones which have recently been shown to be much more bioavailable than previously envisaged [14, 15]

It is becoming increasingly evident that (poly)phenols may also contribute to health benefits through interactions within the gastrointestinal (GI) tract, including protection against epithelial damage associated with GI cancers [16, 17], maintenance of a beneficial microbiota and its associated biotransformation of (poly)phenolic metabolites [18] and inhibition of digestive enzymes that can slow nutrient release with benefits for glycemic control and/or obesity [19-21].

Investigations that simulate digestive processes in vitro [e.g. 22] have provided evidence that berry (poly)phenols have different stabilities in the GI tract and may be available in amounts that could beneficially influence important physiological processes. Although these studies can define and compare potential in-gut availability, they cannot adequately mimic the dynamic, active processes of digestion [23]. Ileostomy studies provide a unique insight and can identify phytochemicals and their metabolites which, in volunteers with an intact colon, would pass from the small to the large intestine [24-26].

Although previous targeted liquid chromatography mass spectrometry (LC-MS^n) analyses [27] confirmed that the major (poly)phenol components of raspberries, anthocyanins and ellagitannins, were present in ileal fluid after raspberry supplementation [26], non-targeted LC-MS^n analyses selected other components that consistently increased in abundance after raspberry supplementation including some previously unreported components [27]. In this study, we sought evidence for the nature of two of the major novel components and extracted and identified these components from raspberries. Through their survival in ileal fluids, these novel components are confirmed as colon-available so we also explored the potential bioactivity of the novel triterpenoids identified using cellular models relevant to colon cancer and discuss the physiological impact of survival in the GI tract of these bioactive components.

2 MATERIALS AND METHODS

2.1 Chemicals

Chemicals were obtained as described previously [27]. Tenuifolin was purchased from Stanford Chemicals Ltd (Irvine CA, USA). 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid (CDDO) was purchased from (Cambridge Bioscience Ltd, UK).

2.2 Plant material and processing

Raspberries (*Rubus idaeus* var. Glen Ample), grown and prepared at the James Hutton Research Institute in 2014 were stored at −80°C. A 8 kg aliquot was defrosted, pureed [7], frozen and transported to University of Ulster for the ileostomy feeding studies.

2.3 Ileostomy feeding study

The ileal fluid samples were collected from the raspberry puree ileostomy feeding study (Ref No. 11/NI/0112) described in full previously [27]. In brief, following a diet low in (poly)phenolic compounds, 11 ileostomates provided a baseline ileal fluid sample (T = 0 h) then consumed 300 g of pureed raspberries and a second ileal fluid sample collected at T= 8 h. The ileal fluid samples were collected, processed within 30 min and stored as aliquots at −80°C.

2.4 Non-targeted LC-MS^n analysis

LC-MS^n analysis of ileal samples was performed on an HPLC system consisting of an Accella 600 quaternary pump, Acella PDA detector coupled to an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) operated under Xcalibur software [27]. The same programme was used to analyse samples from the raspberries.

Certain samples (e.g. seed and puree extracts) were also analyzed using an LCQ-DECA system with an ion trap mass spectrometer [27].

2.5 MS data handling and analysis

The non-targeted LC-MS data from the Orbitrap analysis was first deconvolved using the SIEVE^{TM} software programme which produces a list of retention time (RT)-*m/z* pairs along with an associated extracted ion chromatogram based on peak area for each sample [28]. The data were analysed statistically [27] and a subset was defined of components whose patterns of abundance were increased after supplementation in all 11 subjects

2.6 Raspberry puree and seed extractions

Frozen raspberries (~500g) were thawed on ice then pureed in a Waring blender (top speed, 3 x 15 s). A portion of the puree was sieved to remove seeds (0.5 mm size) and the unseeded puree collected. The seeds were washed with ice-cold distilled water then dried on paper towels. Seeded and unseeded purees (5 mL) and seeds (1 g fresh weight) were extracted with 10 mL of 0.1 % aqueous formic acid for 30 min at 5 ºC with end-over-end mixing in a blood rotator then centrifuged (2 500g, 5 min 5 ºC) and the extracts removed to new tubes. The extraction was repeated with 0.1% aqueous formic acid then sequential extractions were carried with 80% acetonitrile containing 0.1% formic acid and finally with 10 mL 50 % aqueous acetone. Aliquots,1mL, of the extracts were dried in a Speed-Vac, re-suspended in 5% aqueous acetonitrile containing 0.2 % formic acid before LC-MS^n analysis.

2.7 Fractionation

Raspberry puree (15 mL) was diluted in an equal volume of 0.2 M HCl, vortexed well and incubated in a blood rotator (100 rpm) for 10 min at 5 ºC. After centrifugation (2500 g, 10 min, 5 ºC), the supernatant was used for fractionation. Strata X-C solid phase extraction

(SPE) units (100 mg/3 mL units; Phenomenex Ltd, Macclesfield, U.K.) were pre-equilibrated with 10 mL methanol then washed with 10 mL water. Extracts (5 mL) were applied to X-cartridges and unbound material was collected. After a wash with 2 x 5 mL of water, the unbound eluates were combined. The units were eluted with 5 mL methanol which was collected. The cartridges were still red as anthocyanins remained bound. The fractions were tested for total phenol content by the Folin method [29] and only the methanol fraction contained appreciable phenolics. Aliquots (1 mL) were dried, re-suspended as above prior to LC-MS^n analysis.

2.8 Bulk seed extraction and purification

Eight kg of raspberries were pureed in 250 g batches and the seeds separated (seed yield ~ 4% w/w). Seeds (250 g) were extracted with 1 L of 0.1% aqueous formic acid for 60 min at 5 ºC with orbital rotation at 90 rpm and the extract obtained by filtering through a glass sinter (porosity 3). The seeds were then extracted twice with 500 mL of 80% ethanol and finally with 500 mL of 50% aqueous acetone. These fractions were assayed for total phenol content and aliquots (1 mL) dried and re-suspended as above for LC-MS^n analysis. The aqueous extract and the first ethanol extract were combined and diluted to 10 % aqueous ethanol then applied to a 70 mL column of Sephadex LH20 (GE Healthcare, Buckinghamshire, UK) which had been equilibrated with 80 % aqueous acetone then 10 % aqueous ethanol. The unbound fraction was collected and the column was washed with 2 column volumes of 10 % ethanol before being eluted with 80% ethanol then 50 % acetone. Aliquots of each fraction were dried using the Speed-Vac prior to LC-MS analysis. The bulk of the unbound and wash fractions were combined and the ethanol removed by rotary evaporation. The fraction was made up to 0.1 % aqueous formic acid and applied to a C18 solid phase extraction units (Strata C18-E, GIGA units, 10 g capacity; Phenomenex, Ltd., Macclesfield, U.K.) that had been treated with 80 % aqueous acetonitrile containing 0.1% formic acid then 0.1% aqueous formic acid. The unbound fraction was collected with a wash fraction of 75 mL 0.1% aqueous

formic acid. Components were eluted with 75 mL of 15 % aqueous acetonitrile containing 0.1% formic acid followed by 75 ml 80 % aqueous acetonitrile containing 0.1% formic acid. Fractions were assayed for total phenol content and aliquots dried and re-suspended for LC-MS analysis. Enrichment in triterpenoids was followed using the red-brown colour reaction of the Lieberman-Burchard method [31]. The content of triterpenoids in the final fraction was estimated as tenuifolin equivalents by peak areas.

2.9 Tissue culture

Human colon cells HT29 (adenocarcinoma) and CCD 841 CoN (normal epithelial) (32) were acquired from European Collection of Cell Cultures (ECACC) and American Type Culture Collection (ATCC) respectively. HT29 cells were cultured in DMEM supplemented with 10% FBS and 100U/l penicillin/streptomycin. CCD 841 CoN cells were maintained in MEM supplemented with 10% FBS, 100U/L penicillin/streptomycin, 1% sodium pyruvate, 1% NEAA and were used between passage 15-25. Both cell lines were incubated at 37°C with 5% CO_2 and grown as monolayers in roux flasks. Cells were sub-cultured every 3-4 days by the addition of trypsin (0.25% trypsin-EDTA) at 37°C for 5min. Cells were centrifuged at 1200 rpm for 3 min, the supernatant decanted and cells re-suspended in the appropriate medium.

For the purposes of this study both cell lines where treated with either a pure synthetic triterpenoid 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid (CDDO), a known inducer of the Nrf2/ARE pathway (33, 34), or a triterpenoid-rich fraction (TRF). Both treatments were prepared fresh in the respective cell media to a final concentration of 100 nM for use in the in vitro experiments.

2.10 Cytotoxicity and genotoxicity of triterpenoids

2.10.1 *Cytotoxicity assay.* The effects of CDDO and TRF on the viability of HT29 and CCD841 cells were determined using the MTT assay (35). The assay is based on the ability of living cells to metabolize the tetrazolium salt, MTT (3-(4,5-dimethylthaizaol-2-yl)-2,5-diphenyl tetrazolium bromide, Promega, Madison, USA) by mitochondrial activity to formazan, a blue dye (36) which can be measured spectrophotometrically. Cells were seeded in 96 multi-well plates (Costar, Cambridge, MA, USA) at a concentration of 1.5×10^4 HT29 and 3.0×10^4 CCD 841 cells per well respectively. After 2 days incubation at 37°C, media was replaced with 100 nM of either TRF or CDDO, then incubated for 24 h. The wells were washed and cells incubated for a further 48 h in fresh media. Thereafter 15 μL of MTT were added to each well. After 4 h, lysis was carried out with 100 μL solubilizing solution to free the product formazan. Formazan was measured using a microtiter plate reader (Alpha, SLT Rainbow Thermo, Antrim, UK) at a wavelength of 560 nm. The survival of the cells treated with cell media only was set as 100% viability. Each treatment was performed in octuple and the experiment was carried out on 3 separate occasions.

2.10.2 COMET assay. The assay, as described previously [37] using the well-established HT29 cell model for colonic DNA damage (17) and the normal colonocyte CCD841. In brief, both cell lines were incubated for 24 h with 100 nM of either TRF or CDDO. To assess the anti-genotoxic potential of the treatments, the cells were treated with hydrogen peroxide (75μM, H_2O_2 for HT29 and 25 μM, H_2O_2 for CCD 841) for 5 min at 4°C, then centrifuged for 5 min at 258 x g. The supernatant was discarded and the cell pellet re-suspended in 85 μL of 0.85% low melting point agarose (LMPA) in PBS and maintained in a water bath at 40°C. The suspension was added to previously prepared gels (1% normal agarose) on frosted slides and coverslips were added. The slides were immersed in lysis buffer (2.5 M NaCl, 100 mM Na_2EDTA, 10mM TRIS, pH10) for 1 h at 4°C and then placed in electrophoresis buffer and allowed to unwind for 20 min before running at 26 V (300 mA) for 20 min. After electrophoresis, gels were washed 3 times for five min in neutralisation buffer (0.4 M Tris, pH

7.5) at 4°C. All gels were stained with 20 µL of ethidium bromide (2 µg/mL in double distilled water prior to scoring. Images were analysed at 400 times magnification using a Nikon eclipse 600 epi-fluorescence microscope. The percentage DNA tail was recorded using Komet 5.0 image analysis software (Kinetic Imaging Ltd, Liverpool, UK). For each slide, 50 cells were scored. Data represents the mean percent tail DNA of triplicate gels per treatment from three independent experiments. To assess the genotoxic potential of the compounds, cells were treated as above omitting the H_2O_2 challenge. Positive (H_2O_2) and negative controls (PBS) were included in all experiments (cells without CDDO/TRF pre-treatment).

2.11 RNA isolation and cDNA synthesis

Cell pellets were collected from both treated HT29 and CCD 841 CoN cells and homogenised using QIAshredder (Qiagen) prior to extraction of total RNA with RNeasy Mini Plus Kit (Qiagen) as per the manufacturer's instructions. Standard PCR of isolated RNA confirmed the absence of genomic contamination. Quantification of RNA was determined via NanoDrop™ ND-100 UV/VIS spectrophotometer with quality verified by gel electrophoresis. Only RNA samples with A_{260}/A_{280} and A_{260}/A_{230} between 1.8–2.0 were deemed acceptable for downstream applications.

cDNA was produced using anchored-oligo$(dT)_{18}$ and Transcriptor First Strand cDNA synthesis Kit (Roche). Using 1 µg of total RNA, 20 µL of cDNA was synthesised as per the manufacturer's instructions for reverse transcription. Standard PCR using 1 µL cDNA as a template and the housekeeping gene GAPDH (primers shown in Table S1) as a control, confirmed the absence of unspecific products.

2.12 Real-time PCR (qPCR)

Real-time qPCR was performed using the Lightcycler 480 II (Roche) in accordance with the manufacturer's instructions. Each 10 µL PCR reaction contained 0.5 µM each primer, 5 µL LightCycler 480 SYBR Green I Master (Roche), 2 µL nuclease-free water and 1 µL cDNA template. Cycling conditions were as follows: 95°C for 10 min, 50 cycles of 95°C for 10 s, 57°C for 10 s and 72°C for 10 s. Melt curve analysis for each gene confirmed product specificity, and only artefact free reactions were considered valid. For all experiments negative control reactions (no template control and negative reverse transcriptase) were run on the same plate.

cDNA standards (in triplicate) were produced by the 5-fold dilution of pooled cDNA, and used to produce cDNA calibration curve slopes. LightCycler software (version 1.5) generated primer efficiencies for each gene (primers shown in Table 1), calculated by using the equation $E=10^{[-1/\text{slope}]}$ [38]. cDNA templates used for target runs were diluted to within the standard curve linear range, in this case 1:10. Only primers with an efficiency of 90-100% were used for target analysis (Table S1)

Relative expression is calculated using the following equation:

$$\frac{Conc.\ target\ gene}{Conc.\ reference\ genes}\ (control) : \frac{Conc.target\ gene}{Conc.reference\ genes}\ (sample)$$

Each cDNA target sample was normalised to 3 reference genes, (HPRT, β-Actin, GAPDH) and calculated as a ratio of the untreated control samples. All target cDNA samples were performed as technical triplicates, with biological replicates for each sample.

2.13 Statistical Analysis

The mean of each data set was used for statistical analysis and experiments were carried out as independent triplicates. The Shapiro-Wilk test was used to test for normality. Analysis of variance was applied to test for significant differences between means compared to

control using Dunnett T post hoc test. Significance was accepted at $p < 0.05$. Analysis was carried out using SPSS (version 20 for Windows).

3 Results

Two major unknown compounds (*m/z* 355 and 679) were consistently identified in ileal fluids from the 11 volunteers [27] after raspberry supplementation (Fig. 1). The compound at *m/z* 679 was present at a similar intensity to the major ellagitannin peak, sanguiin H-6 (compare Figs. 1A and 1B). The compound at *m/z* 355 (Fig. 1C) was present in three separate peaks (see Supplementary data; Fig. S1). Initial $[M-H]^-$ MS data (Table 1) [*m/z* = 355.0648 and MS^2 fragments of 337, 209 and 191] did not match with any berry component or database entry [27]. However, positive mode data gave *m/z* M+H = 357.0809, predicted formula of $C_{15}H_{17}O_{10}$ with MS^2 of *m/z* 339 (loss of H_2O) and 147 (loss of 210). In-source fragments mirrored the MS^2 fragments and gave predicted formulae for *m/z* $M+H^+$ = 339 of $C_{15}H_{15}O_9$ and *m/z* M+H = 147 of $C_9H_7O_2$. The neutral loss of 146 amu in negative mode and the MS^2 fragment at 147 in positive mode suggests a *p*-coumaric acid derivative. Indeed, the peak had an absorbance maximum at ~310 nm, characteristic of hydroxycinnamates (results not shown). The neutral loss of 210 in positive mode and the MS^2 fragment at *m/z* 209 in negative mode can be assigned to glucaric acid which strongly suggests that these components are *p*-coumaroyl glucarate isomers (see diagram 1, structure 1)) as described in dog's mercury [39]. These compounds have not been previously identified in raspberry but caffeoyl glucarates have been identified in calafate berries [40]. *p*-Coumaroyl glucarate components were confirmed in extracts of Glen Ample raspberries (see Supplementary data; Fig. S2A & B) and were enriched by fractionation on strong ion-exchange SPE cartridges which retained anthocyanins and ellagitannins but released the coumaroyl glucarates (Supplementary data; Fig. S2C & D).

The *m/z* 679 signal gave no UV absorbance above 240 nm and so was unlikely to be a (poly)phenolic compound. Exact mass of 679.3648 yielded a predicted formula of $C_{36}H_{55}O_{12}$ (Table 1) and the major MS^2 fragment at 517 (loss of 162) suggested a glycoside with an aglycone of 517. However, the *m/z* 679 signal had 0.5 amu variants (Table 1) which suggests a doubly charged entity with a true mass of ~1360. A signal at *m/z* at 1359.7403 was present (predicted formula $C_{72}H_{111}O_{24}$). Positive mode data confirmed these results [exact mass 681.3833; predicted formula of $C_{36}H_{57}O_{12}$; exact mass 1361.7603, predicted formula of $C_{72}H_{113}O_{24}$). However, no MS^2 of the M+H 1361.7 or M-H 1359.7 were obtained. Compounds which were good matches for the predicted formula obtained for $[M-H]^-$ *m/z* 679 (MW $C_{36}H_{56}O_{12}$) and were glycosides were all isomers of ursolic acid-based triterpenoid glucosides that differed in the position of attachment of glucose, hydroxyl or methyl groups etc. (Table 1; Diagram 1; structure 2). Triterpenoids have been identified in raspberry leaves and flowers, including in commercial varieties of *Rubus ideaus* [41]. Ursolic acid based triterpenoid glycosides have been noted in leaves of *R. coreanus* [e.g. suavissimoside R1 [42], whereas trachelosperoside B1 (MW = 682) and nigaichigoside F1 (MW = 666) have also been identified in *Rubus rosifolius* fruits [43]. Ilexoside XLV has been reported to give a main MS^2 fragment at 517 [44] from loss of hexose but any of the isomers reported in Table 1 could give the same fragmentation.

Two separate peaks (T1 and T2) with apparent *m/z* 679 were apparent in methanol extracts of seeds and were major contributors to the total MS signal (Fig. 2). Only peak T1 was detected in the ileal fluids. Sequential extractions of raspberry puree, isolated seeds and "unseeded" puree (see supplementary material Fig. S3) showed that Peaks T1 and T2 were present in both whole and unseeded purees but with higher levels in the whole puree (Fig. 3) and seeds. Peak T2 was more abundant in the acetonitrile and acetone extracts than water extracts suggesting that it was more hydrophobic than peak T1. Therefore, although the MS data suggests that peaks T1 and T2 may be related to the ursolic acid triterpenoid

glycosides noted in Table 1, the nature of these components was examined further after purification.

Fractionation of seed extracts removed polyphenols and enriched peaks T1 and T2 (Fig. 4). This enrichment was accompanied by reduced total phenol content, enhanced response to the Liebermann–Burchard reaction for triterpenoids (results not shown) and an enrichment of other putative triterpenoid peaks (Fig. 4, Table 2; peaks 13, 14, A & B). Indeed, alkaline hydrolysis of the initial seed extract produced *m/z* signals from simple phenolics from degradation of the polyphenols and a range of triterpenoid aglycones consistent with peaks T1 and T2 (i.e. *m/z* 517) but also other putative triterpenoids (see supplementary data Fig. S4, Table S2 for data and method).

The MS properties of peak T1 were similar to those of the *m/z* 679 signal in ileal samples having 0.5 amu variants suggesting doubly charged status and an actual mass of ~1360 (Fig. 5; Table 2). The signal at *m/z* 1359.7643 (predicted formula $C_{72}H_{111}O_{24}$) gave a single MS^2 fragment at 679.3. Therefore, Peak T1 may be similar to the ester-linked dimeric triterpenoid, Coreanoside F1 ($C_{72}H_{110}O_{24,}$ MW = 1358.7; Diagram 1; structure 3) extracted from *Rubus coreanus* leaves [45], but this compound has an MW 2 amu less than the apparent MW for peak T1. This "extra 2H" could occur by (e.g.) substitution with $–CH_2OH$ and –COOH groups at one position (+14 amu) and –OH and -H groups at another position (-16; net difference +2 amu).

Peak T2 had a major *m/z* at 1357.7 (Fig. 5; Table 2), by analogy with *m/z* 1359 in peak T1, this could result from a component with $C_{72}H_{109}O_{24}$ but this was not predicted from the accurate mass data. The signal at *m/z* 1357.7 gave no MS^2 but the signal at 1358.7 yielded a major MS^2 fragment at 679.4 (loss of 679) with fragments at 1313.4 (loss of 46); 1196.4 (loss of 162), 1151.7 (loss of 162 & 46) 1018.7 (loss of 340) and also 559.7 and 517.3. Considering that peak T2 also showed 0.5 amu variants suggesting doubly charged status and an nominal actual mass of ~ 2716, the MS data strongly suggests that peak T1 and T2 are structurally related, and that peak T2 could be a dimer of peak T1. Initial ^1H NMR spectra

were consistent with peak T1 being a ursolic acid based triterpenoid but further NMR studies and use of alternative MALDI-TOF MS techniques with higher *m/z* ranges [46] will be required to confirm the nature of these putative triterpenoid derivatives.

After purification, the triterpenoid-rich fraction (TRF) was effective in preventing H_2O_2-induced DNA damage to HT29 adenocarcinoma cells and also to the normal epithelial colon cell line CCD 841 CoN (Fig. 5). At 100 nM, both TRF and the synthetic triterpenoid 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid (CDDO) exerted a significant anti-genotoxic effect against H_2O_2 challenge in the two cell lines (Fig. 5). In HT 29 cells, both treatments reduced tail DNA by ~40-45 % compared to the untreated control, whilst in CCD 841 CoN cells, DNA damage was reduced by 50-55%. Efficacy between cell lines was not significantly different. No cytotoxic or genotoxic activity was observed for either TRF or CDDO at 100 nM (data not shown).

CDDO has been reported to induce the Nrf2/ARE pathway including NAD(P)H dehydrogenase, quinone-1 (NQO1) and heme oxygenase-1 (HO-1) [33]. These enzymes reduce reactive oxygen species and play a key role in cytoprotection; NQO1 acts a superoxide scavenger [47] and HO-1 reduces the production of free radicals via the catabolism of heme [48]. After 24 h exposure, CDDO significantly increased expression of both NQO-1 and HO-1 in HT29 and CCD 841 CoN cells (Fig. 6), but Nrf2 expression was significantly reduced. The 24 h time-frame was chosen to match with the genoprotective studies but a time-course experiment (see Fig S5) showed that CDDO enhanced Nrf2 expression in both cell lines over shorter exposure times.

TRF increased the expression of the Nrf2 gene in HT29 cells but reduced expression in normal CCD 841 CoN cells (Fig. 6) following 24hr exposure. TRF showed a small, but significant, increase in NQO1 expression in CCD 841 cells but HO-1 expression was significantly reduced, in both HT29 and CCD841 cells. This differential effect on activation of the Nrf2-regulated pathway may be caused by the concentration difference between the pure CDDO and TRF (which is a mixture of triterpenoids), or it may reflect structural

differences between the compounds. Indeed, the position of methyl groups in the triterpenic dialcohols, uvaol and erythrodiol, from olive oil altered bioactivity from genoprotective to genotoxic in both normal and breast cancer cell lines (MCF10A & MDA-MB-231) [49]. Alternatively, it could represent a time-dependent effect on up-regulation of Nrf2 expression in both cells as suggested by the results with CDDO (Figure S5).

4 Discussion

The application of non-targeted LC-MS^n analysis to ileal fluids after raspberry intake selected previously unknown components that arose from the berry intake. Further extraction and purification work confirmed their presence in raspberries and allowed their identification as *p*-coumaroyl glucarates and ursolic acid-based triterpenoid glucosides. Their relative stability to GIT conditions explains their revelation in ileal fluid and confirms their bioavailability throughout the GIT and into the colon. However, the non-targeted nature of the LC-MS based approach and experimental robustness of the data from the ileostomy cohort was essential in bringing these previously unknown components to light.

The presence of *p*-coumaroyl glucarate derivatives in raspberries, and their survival in ileal fluids, could be of significance as D-glucaric acid has long been known to have anti-cancer effects in carcinogen-induced animal models [50]. Although originally thought to act through inhibition of β-glucuronidase and increased detoxification of carcinogens, it may have more complex effects via anti-inflammatory systems [51] and apoptotic events [52]. As these coumaroyl glucarate esters survive into the ileal fluid, they could deliver potentially-bioactive glucaric acid throughout the GI tract, ultimately to the colon, where it could have a role in chemoprevention of cancer. Future work will purify and assess the potential bioactivity of these colon-available raspberry components.

The fraction from raspberries enriched in the novel triterpenoids was found to be effective in protecting against H_2O_2-induced DNA damage at sub-micromolar levels, which

could be readily achieved and therefore physiologically relevant in vivo. Indeed, the parent triterpenoid ursolic acid decreased H_2O_2-induced DNA damage in the colonic cell line Caco-2 by a similar extent (>40%) to that of TRF and CDDO, albeit at 50-fold higher concentration of 5 μM [53]. Triterpenoids from *Rubus rosifolius* fruits have previously been shown to have beneficial effects against human colon cancer cells [42] and triterpenoids from *Scoparia dulcis* roots exerted anti-mutagenic activity using the *in vivo* SMART assay [54]. To our knowledge, this study is the first to demonstrate DNA damage reduction and the modulation of Nrf2/ARE pathway by triterpenoids (both CDDO and TRF) in the normal cell line CCD841-CoN. Although recent work [55] reported that sulforaphane decreased H_2O_2-mediated oxidative damage and activated the Nrf2/ARE pathway in CCD841 cells, concentrations of 2.5-40μM were required to elicit this effect. Triterpenoids have been reported to have other relevant bioactivities; oleanolic acid glycosides influence gastrointestinal transit in mice [56]; triterpenoids modulate intestinal transport [57]; purified triterpenoids from *Rubus parvifolius,* including suavissimoside R1 and coreanoside F1, have anti-fatigue effects in mice [58] and triterpenoid-rich fractions from Korean raspberries have potent anti-inflammatory effects [59]. In many cases, these triterpenoids are isolated from non-edible plant parts such as roots or leaves, which would not form part of the normal diet.

It is possible that these raspberry fruit components may have contributed to the positive effects noted for raspberry extracts in our previous studies on colon cancer models [17, 22] and may have contributed to bioactivities assigned to berries in other studies [14]. Indeed, after re-examination of previous data we can confirm that the putative triterpenoids survived in vitro digestion procedures used to simulate gastrointestinal conditions (e.g. [22, 37].

In conclusion, these novel components were discovered by their survival in ileal fluids, they are available in the gut and we can assume that they would enter the colon in vivo and could exert these bioactivities in situ throughout the GI tract. Further work on the contribution of these novel components to potential health effects of raspberries is merited.

*The authors responsibilities were as follows***:** *CG, GMcD, RL, IR and AC were involved in study design and CG, and GO'C in study conduct. GMcD, SV, CL, MI, GP, CG, EMB and JWA were involved in experimental and data analysis. The manuscript was prepared by GMcD, CG, IR, AC, DS, GMcM and JWA.*

We would like to thank the volunteers for participating in the study. CG, RL and AC acknowledge funding from the National Processed Raspberry Council. GMcD, SV, DS and JWA acknowledge funding from the Scottish Government's Rural and Environment Science and Analytical Services (RESAS) Division. DS and GMcD acknowledge funding from BachBerry (Project No. FP7-613793).GPC was supported by a postdoctoral fellowship from IFAPA (Programa Operativo del Fondo Social Europeo 2007−2013 de Andalucía).

The authors declare no conflicts of interest.

5 References

[1] Hartley, L., Igbinedion, E., Holmes, J., Flowers, N. et al., Increased consumption of fruit and vegetables for the primary prevention of cardiovascular diseases. *Cochrane Database Syst. Rev*. 2013, 6,CD009874. doi: 10.1002/14651858.CD009874.pub2.

[2] Ezzati, M., Riboli, E. GLOBAL HEALTH: Behavioural and dietary risk factors for non-communicable diseases. *New England J. Med*. 2013, *369*, 954–964.

[3] Hjartåker, A., Knudsen, M. D., Tretli, S., Weiderpass, E., Consumption of berries, fruits and vegetables and mortality among 10,000 Norwegian men followed for four decades. *Eur. J. Nutr*. 2015, *54*: 599–608.

[4] Del Rio, D., Rodriguez-Mateos, A., Spencer, J. P., Tognolini, M. et al., Dietary (poly)phenolics in human health: structures, bioavailability, and evidence of protective effects against chronic diseases. *Antioxid. Redox Signal*. 2013, *18*, 1818–1892.

[5] Li, A. N., Li, S., Zhang, Y.J., Xu, X.R., Chen, Y.M., Li, H.B., Resources and biological activities of natural polyphenols. *Nutrients* 2014, *6*, 6020–6047.

[6] Rodrigues-Mateos, A. M., Vauzour, D., Kreuger, C. G., Shanmuganayagam, D. et al., Flavonoids and related compounds, bioavailability bioactivity and impact on human health: An update. *Arch. Toxicol*. 2014, *88*, 1803–1853.

[7] Joseph, S. V., Edirisinghe, I., Burton-Freeman B. M., Berries: Anti-inflammatory effects in humans. J. Agric. *Food Chem*. 2014, *62*, 3886–3903.

[8] Seeram, N. P., Berries and human health: Research highlights from the Fifth Biennial Berry Health Benefits Symposium. *J. Agric. Food Chem*. 2014, *62*, 3839–3841.

[9] Rodriguez-Mateos, A. M., Heiss, C., Borges, G., Crozier, A. (2014) Cardioprotective effects of berry (poly)phenols. *J. Agric. Food Chem.* 2014, 62, 3842–3851

[10] Prior, R. L., Anthocyanins: understanding their absorption and metabolism. In "*Flavonoids and Related Compounds. Bioavailability and Function*", (Spencer, J. P. E.; Crozier, A. eds.) CRC Press, Boca Raton, FL, pp. 79–92; 2012.

[11]. Czank, C., Cassidy, A., Zhang, Q., Morrison, D. et al., Human metabolism and elimination of the anthocyanin, cyanidin-3-glucoside: A ^{13}C-tracer study. *Am. J. Clin. Nutr.* 2013, *97*, 995–1003.

[12] de Ferrars, R. M., Czank, C.; Zhang, Q.; Botting, N. P. et al., The pharmacokinetics of anthocyanins and their metabolites in humans. *Br. J. Pharmacol.* 2014, *171*, 3268–3282.

[13] Ludwig, I. A., Mena, P., Calani, L., Borges, G. et al., New insights into the bioavailability of red raspberry anthocyanins and ellagitannins. *Free Radic. Biol. Med.*, 2015, *89*, 758–769.

[14] Pereira-Caro, G., Borges, G., van der Hooft, J., Clifford, M. N., Lean, M. E. J., et al. Orange juice (poly)phenols are highly bioavailable. *Am. J. Clin. Nutr.* 2014, *100*, 1385–1391

[15] Pereira-Caro, G., Oliver, C.M., Weerakkody, R., Singh, T., Conlon, M. et al. Chronic administration of a microencapsulated probiotic enhances bioavailability of orange juice flavanones in humans. *Free Radic. Biol. Med.* 2015, *84*, 206–214

[16] Petrick, J. L., Steck, S. E., Bradshaw, P. T., Cairney, F. et al., Dietary intake of flavonoids and oesophageal and gastric cancer: incidence and survival in the United States of America (USA). *Br. J. Cancer*. 2015, *112*, 1291–1300.

[17] Brown, E. M., Latimer, C., Allsopp, P., Ternan, N. G. et al., *In vitro* and *in vivo* models of colorectal cancer: Anti-genotoxic activity of berries. *J. Agric. Food Chem.* 2014, *62*, 3852–3866.

[18]. Marín, L., Miguélez, E. M., Villar, C. J., Lombó, F., Bioavailability of dietary polyphenols and gut microbiota metabolism: Antimicrobial properties. *Biomed. Res. Int*. 2015, doi: 10.1155/2015/905215.

[19] Burton-Freeman, B., Postprandial metabolic events and fruit derived phenolics: A review of the science. *Br. J. Nutr*., 2010, *104, S3*, S1–14.

[20] Boath, A., Grussu, D., Stewart, D., McDougall, G. J., Berry polyphenols inhibit digestive enzymes: A source of potential health benefits? *Food Digestion*, 2012, *3*, 1–7.

[21] Hoggard, N., Cruickshank, M., Moar, K. M., Bestwick, C. et al., A single supplement of a standardised bilberry (*Vaccinium myrtillus* L.) extract (36 % anthocyanins) modifies glycaemic response in individuals with type 2 diabetes controlled by diet and lifestyle. *J. Nutr. Sci.* 2013, *2*, e22.

[22] Coates, E. M., Popa, G., Gill, C. I., McCann, M. J. et al., Colon-available raspberry polyphenols exhibit anti-cancer effects on *in vitro* models of colon cancer. *J. Carcinogen*. 2007, *18*, 1–6.

[23] Bohn, T., McDougall, G. J., Alegría, A., Alminger, M. et al., Mind the gap - deficits in our knowledge of aspects impacting the bioavailability of phytochemicals and their metabolites: A position paper focusing on carotenoids and polyphenols. *Mol. Nutr Food Res.* 2015, *59*, 1307–1323.

[24] Kahle, K., Kraus, M., Scheppach, W., Ackermann, M. et al., Studies on apple and blueberry fruit constituents: do the polyphenols reach the colon after ingestion? *Mol. Nutr. Food Res*. 2006, *50*, 418–23.

[25] Borges, G., Mullen, W., Mullan, A., Lean, et al., Bioavailability of multiple components following acute ingestion of a polyphenol-rich juice drink. *Mol. Nutr. Food Res*. 2010, *54*, S268–277.

[27] González-Barrio, R., Borges, G., Mullen, W., Crozier, A., Bioavailability of anthocyanins and ellagitannins following consumption of raspberries by healthy humans and subjects with an ileostomy. *J. Agric. Food Chem*. 2010, *58*, 3933–3939.

[27] McDougall, G. J., Conner, S., Pereira-Caro, G., Gonzalez-Barrio, R. et al., Tracking (poly)phenol components from raspberries in ileal fluid. *J. Agric. Food Chem.* 2014, *62*, 7631−7641.

[28] http://www.thermoscientific.com/en/product/sieve-software-differential-expression.html.

[29] He, J., Giusti, M. M., High-purity isolation of anthocyanins mixtures from fruits and vegetables - a novel solid-phase extraction method using mixed mode cation-exchange chromatography. *J Chromatogr. A.* 2011, *1218*, 7914–7922.

[30] Deighton, N., Brennan, R., Finn, C., Davies, H. V., Antioxidant properties of domesticated and wild Rubus species. *J. Sci. Food Agric*. 2000, *80*, 1307–1313.

[31] Burke, R.W., Diamondstone, B. I., Velapoldi, R. A., Menis, O. Mechanisms of the Liebermann-Burchard and Zak Color Reactions for Cholesterol. *Clin. Chem*. 1974, *20*, 794–801.

[32] Kan, W. L., Yin, C., Xu, H. X., Xu, G., et al., Antitumor effects of novel compound, guttiferone K, on colon cancer by p21Waf1/Cip1-mediated G(0) /G(1) cell cycle arrest and apoptosis. *Int. J. Cancer*. 2013, 132, 707–716.

[33] Liby, K., Hock, T., Yore, M. M., Suh, N. et al., The synthetic triterpenoids, CDDO and CDDO-imidazolide, are potent inducers of heme oxygenase-1 and Nrf2/ARE signaling. *Cancer Res*. 2005, *65*, 4789-4798.

[34] Bensasson, R., V., Zoete, V., Berthier, G., Talalay, P., Dinkova-Kostova, A., T. Potency ranking of triterpenoids as inducers of a cytoprotective enzyme and as inhibitors of a cellular inflammatory response via their electron affinity and their electrophilicity index. *Chem. Biol. Interact*. 2010,*186*, 118–126.

[35] Hansen, M. B., Nielsen, S. E. and Berg, K. Re-Examination and further development of a precise and rapid dye method for measuring cell-growth/cell kill. *J. Immunol. Methods* 1989, *119*, 203–210.

[36] Mosmann T: Rapid colorimetric assay for cellular growth and survival - application to proliferation and cytotoxicity assays. *J. Immunol. Methods 65*, 55–63, 1983.

[37] Brown, E. M., McDougall, G. J., Stewart, D., Pereira-Caro, G. et al., Persistence of anticancer activity in berry extracts after simulated gastrointestinal digestion and colonic fermentation. *PloS One*. 2012, *7*, e49740.

[38] Pfaffl, M. W., A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001, *29*, e45.

[39] Lorenz, P., Conrad, J., Bertrams, J., Berger, M. et al., Investigations into the phenolic constituents of dog's mercury (*Mercurialis perennis* L.) by LC-MS/MS and GC-MS analyses. *Phytochem. Anal*. 2012, *23*, 60–71.

[40] Ruiz, A., Mardones, C., Vergara, C., von Baer, D. et al., Isolation and structural elucidation of anthocyanidin 3,7-β-*O*-diglucosides and caffeoyl-glucaric acids from calafate berries. *J. Agric. Food Chem*. 2014, *62*, 6918–6925.

[41] Griffiths, D. W., Robertson, G. W., Shepherd, T., Birch A. N. et al., A comparison of the composition of epicuticular wax from red raspberry (*Rubus idaeus* L.) and hawthorn (*Crataegus monogyna* Jacq.) flowers. *Phytochem*. 2000, *55*, 111–116.

[42] Zhou, X-H., Kasai, R., Ohtani, K., Tanaka, O. et al., Oleanane and ursane glucosides from *Rubus* species. *Phytochem*istry 1992, *31*, 3642–3644.

[43] Bowen-Forbes, C. S., Mulabagal, V., Liu, Y., Nair, M. G., Ursolic acid analogues: non-phenolic functional food components in Jamaican raspberry fruits. *Food Chem.* 2009, *116*, 633–637.

[44] Amimoto, K., Yoshikawa, K., Arihara, S., Triterpenoid saponins of aquifoliaceous plants. XI. Ilexosides XLI-XLV from the leaves of *Ilex rotunda* Thunb. *Chem. Pharm. Bull.* 1993, *41*, 39–42.

[45] Ohtani, K., Miyajima, C., Takahasi, T., Kasai, R. et al., A dimeric triterpene-glycoside from Rubus. *Phytochemistry* 1990, *29*, 3275–3280.

[46] Hager, T.J., Howard, L. R., Liyanage, R., Lay, J. O., Prior, R. L., Ellagitannin composition of blackberry as determined by HPLC-ESI-MS and MALDI-TOF-MS. *J. Agric. Food Chem.* 2008, *56*, 661–669.

[47] Siegel, D., Gustafson, D. L., Dehn, D. L., Han, J. Y., Boonchoong, P., Berliner, L. J., Ross D. NAD(P)H:quinone oxidoreductase 1: role as a superoxide scavenger. *Mol. Pharmacol*. 2004, 65, 1238–1247.

[48] Gozzelino, R., Jeney, V., Soares, M. P. Mechanisms of cell protection by heme oxygenase-1. Ann. Rev. *Pharmacol. Toxicol*. 2010; 50, 323–354.

[49] Sánchez-Quesada, C., López-Biedma, A., Gaforio, J. J., The differential localization of a methyl group confers a different anti-breast cancer activity to two triterpenes present in olives. *Food Funct*. 2015, *6*, 249–256.

[50] Walaszek, Z., Szemraj, J., Narog, M., Adams, A. K. et al., Metabolism, uptake, and excretion of a D-glucaric acid salt and its potential use in cancer prevention. *Can. Det. Prev*. 1997, *21*, 178–190.

[51] Zoltaszek, R., Kowalczyk, P., Kowalczyk, M. C., Hanausek, M. et al., Dietary D-glucarate effects on the biomarkers of inflammation during early post-initiation stages of benzo[a]pyrene-induced lung tumorigenesis in A/J mice. *Oncology Lett*. 2011, *2*, 145–154

[52] Tiwari, P., Sahay, S., Pandey, M., Qadri, S. S., Gupta, K. P., Combinatorial chemopreventive effect of butyric acid, nicotinamide and calcium glucarate against the 7,12-dimethylbenz(a)anthracene induced mouse skin tumorigenesis attained by enhancing the induction of intrinsic apoptotic events. *Chemico-Biol. Interact*. 2015, *226*, 1–11.

[53] Ramos, A. A., Pereira-Wilson, C., Collins, A. R., Protective effects of ursolic acid and luteolin against oxidative DNA damage include enhancement of DNA repair in Caco-2 cells. *Mutat. Res.* 2010, *692*, 6–11.

[54] de Freitas, P. L., Dias, A. C., Moreira, V. R., Monteiro, S. G., Pereira, S. R., Anti-mutagenic action of the triterpene betulinic acid isolated from *Scoparia dulcis* (Scrophulariaceae*). Genet. Mol. Res*. 2015, *14*, 9745–9752.

[55] Wang, Y., Dacosta, C., Wang, W., Zhou, Z. et al., Synergy between sulforaphane and selenium in protection against oxidative damage in colonic CCD841 cells. *Nutr. Res*. 2015, *35*, 610–617.

[56] Li, Y., Matsuda, H., Yoshikawa, M. Effects of oleanolic acid glycosides on gastrointestinal transit and ileus in mice. *Bioorg. Med. Chem.* 1999, *7*, 1201–1205.

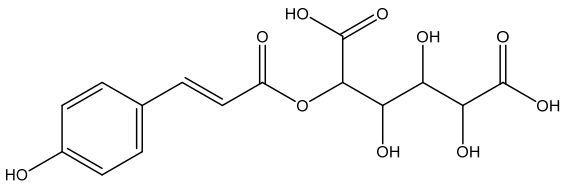
[57] Johnson, I. T., Gee, J. M., Price, K., Curl, C., Fenwick, G. R. Influence of saponins on gut permeability and active nutrient transport *in vitro*. *J. Nutr.* 1986, *116*, 2270–2277.

[58] Chen, J., Wang, X., Cai, Y., Tang, M. et al., Bioactivity-guided fractionation of physical fatigue-attenuating components from *Rubus parvifolius* L. *Molecules* 2013, *18*, 11624–11638.

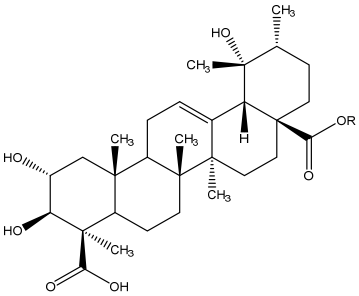
[59] Shin, J-S., Cho, E-J., Choi, H-E., Seo, J-H., An, H-J., Park, H-J., Cho, Y-W., Lee, K-T. Anti-inflammatory effect of a standardized triterpenoid-rich fraction isolated from *Rubus coreanus* on dextran sodium sulfate-induced acute colitis in mice and LPS-induced macrophages. *J. Ethnopharm.*, 2014, *158*, 291–300.

[60] Perret, C., Wolfender, J.-L., K. Hostettmann LC/ES-MS analysis of triterpene glycosides: Rapid estimation of the saponin content of dried berries of *Phytolacca dodecandra*. *Phytochem. Anal.* 1999, *10*, 272–278.

Diagram 1

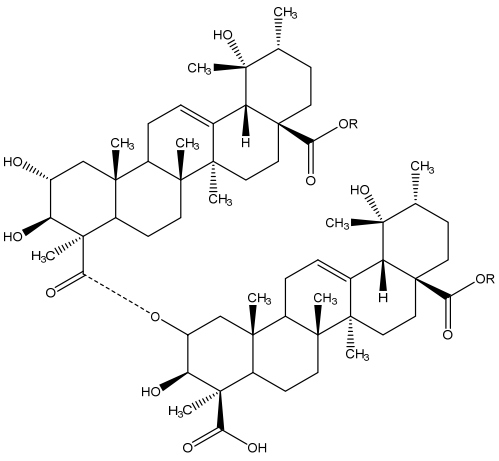


1. *p*-coumaroyl glucarate



2. Trachelosperoside A1 ([1-O-[(2α,3β,5ξ,9ξ)-2,3,19,24-tetrahydroxy-24,28-dioxours-12-en-28-yl]-β-D-glucopyranose]); R = glucosyl.

(http://www.chemspider.com/Chemical-Structure.10273095.html)



3. Coreanoside F1 (https://pubchem.ncbi.nlm.nih.gov/compound/44202896), R = glucosyl

Legends

Figure 1. Comparison of abundance of selected MS signals in ileal fluids before and after raspberry intake. Panel A = *m/z* 934 = sanguiin H-6 at RT 17.9 min; B = *m/z* 679 at RT 23.5; C = *m/z* 355 at RT 12.7. Peak areas are in arbitrary MS units.

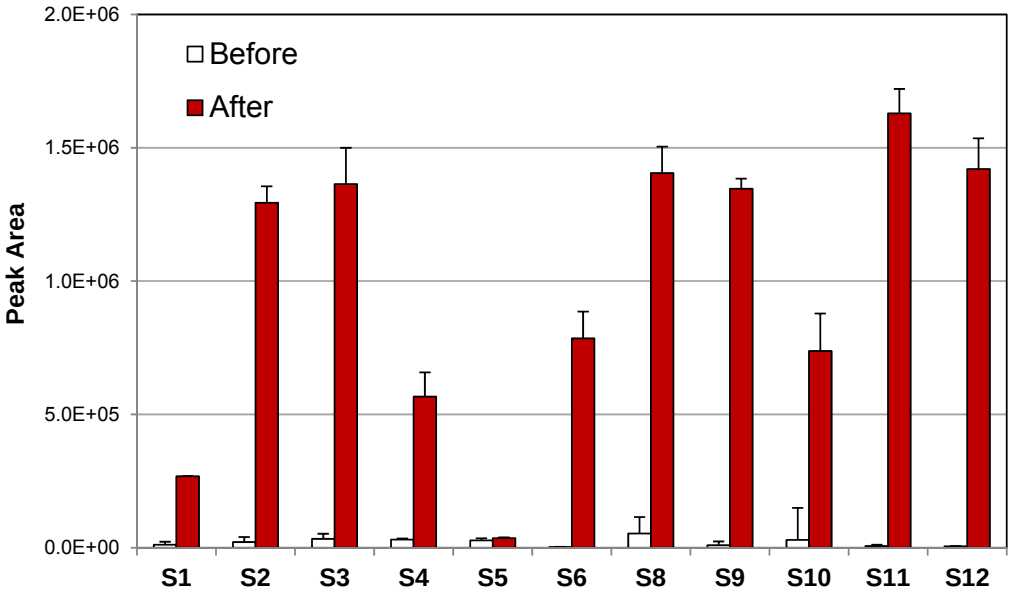
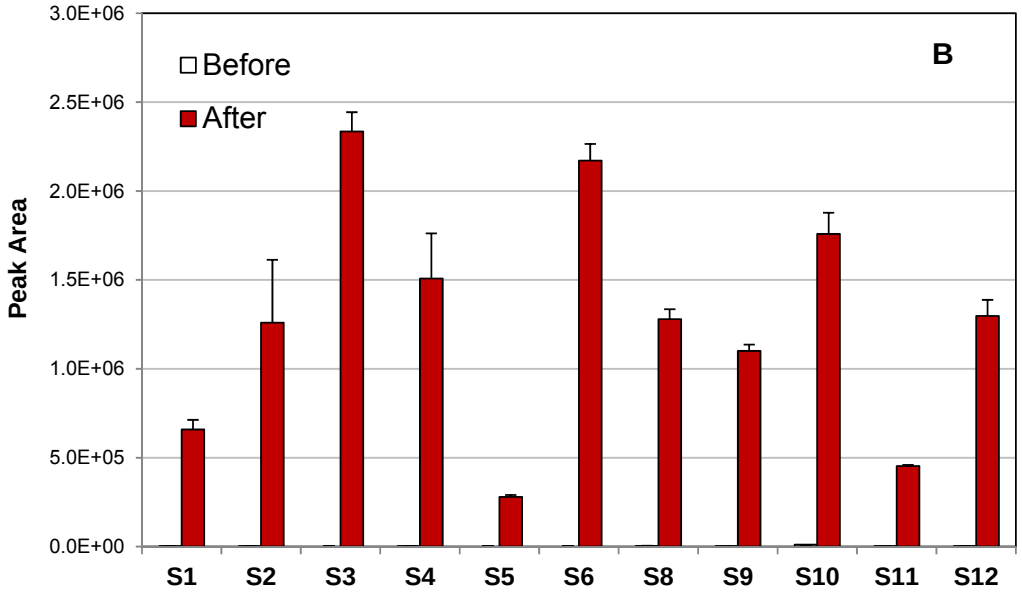


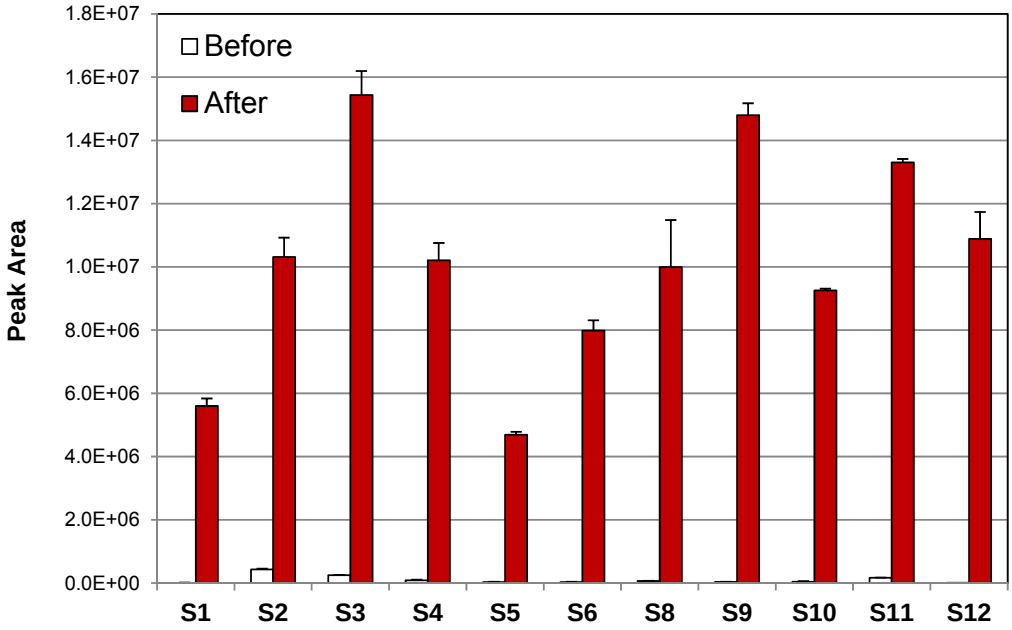
Figure 2. Mass spectral characteristics of *m/z* 679 signal in raspberry seed extracts. Panel A = UV trace at 280 nm; B = MS base peak; C = MS spectra of peak T1; D = MS spectra of peak T2. MS spectra were obtained on ion-trap MS which does not detect doubly charged ions. Figures in the top right corners represent the full scale deflection of the MS detector.

Sanguiin H6; *m/z* 934, RT = 17.9



***m/z* 679, RT = 23.5**

Figure 3. Comparative abundance of peaks T1 and T2 in sequential extracts from whole puree (WP), unseeded puree (USP) and seeds. Peak areas are average of three determinations ± SE. Areas are in arbitrary MS units.



***m/z* 355 main peak, RT = 12.7**

Figure 4. LC-MS profiles of fractions from purification scheme. Panel A = 80 % ethanol seed extract; B = LH20 Sephadex -unbound fraction, C = LH20 Sephadex 10 % ethanol wash fraction, D = 80% ACN SPE fraction. Figures in the top right corners represent the full-scale deflection of the MS detector.

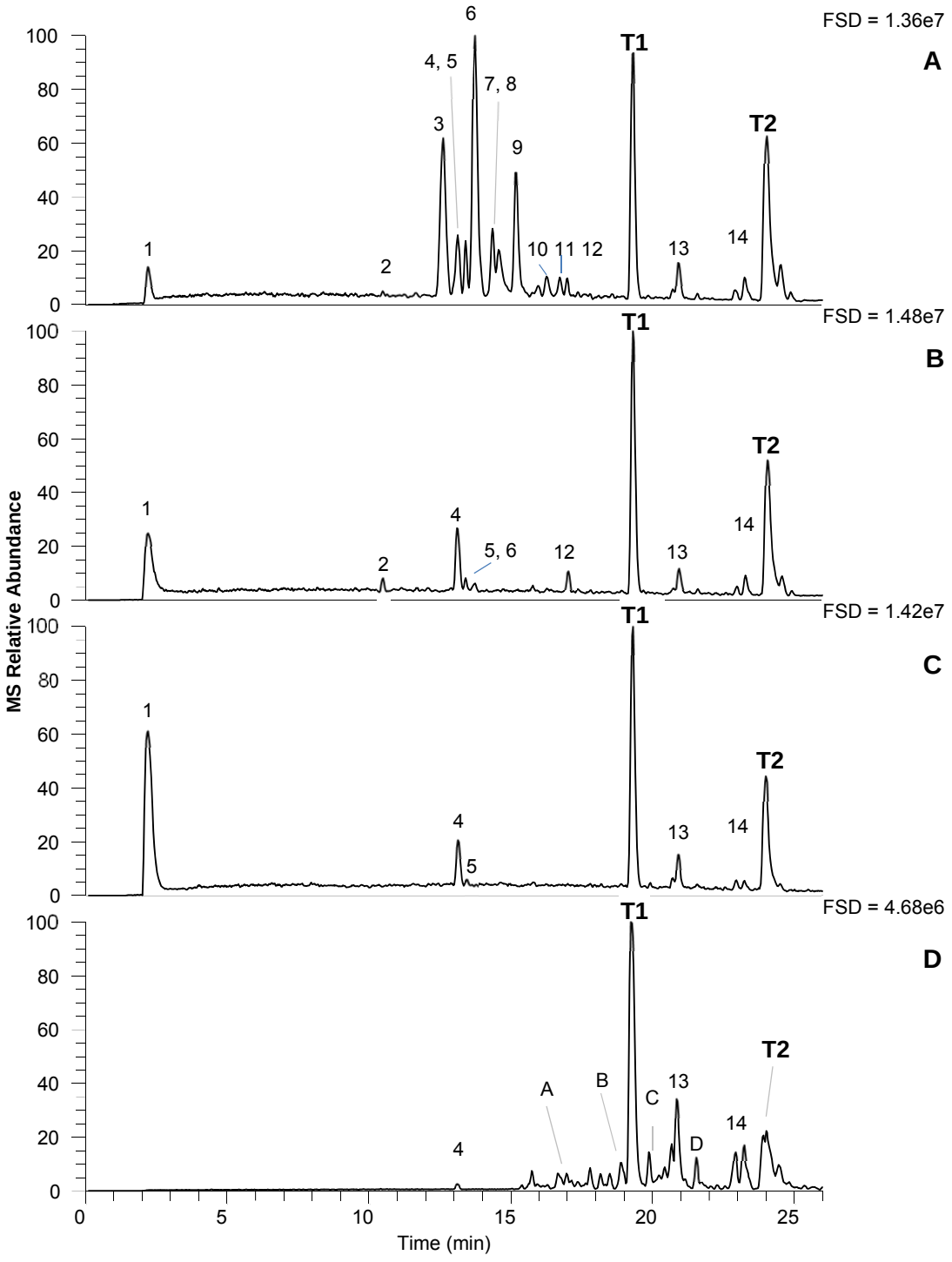


Figure 5. MS spectra of LC peaks T1 and T2. Panel A = peak T1, B = peak T2. *m/z* values in bold are discussed in the text and in Table 2. Figures in the top right corners represent the full-scale deflection of the MS detector.

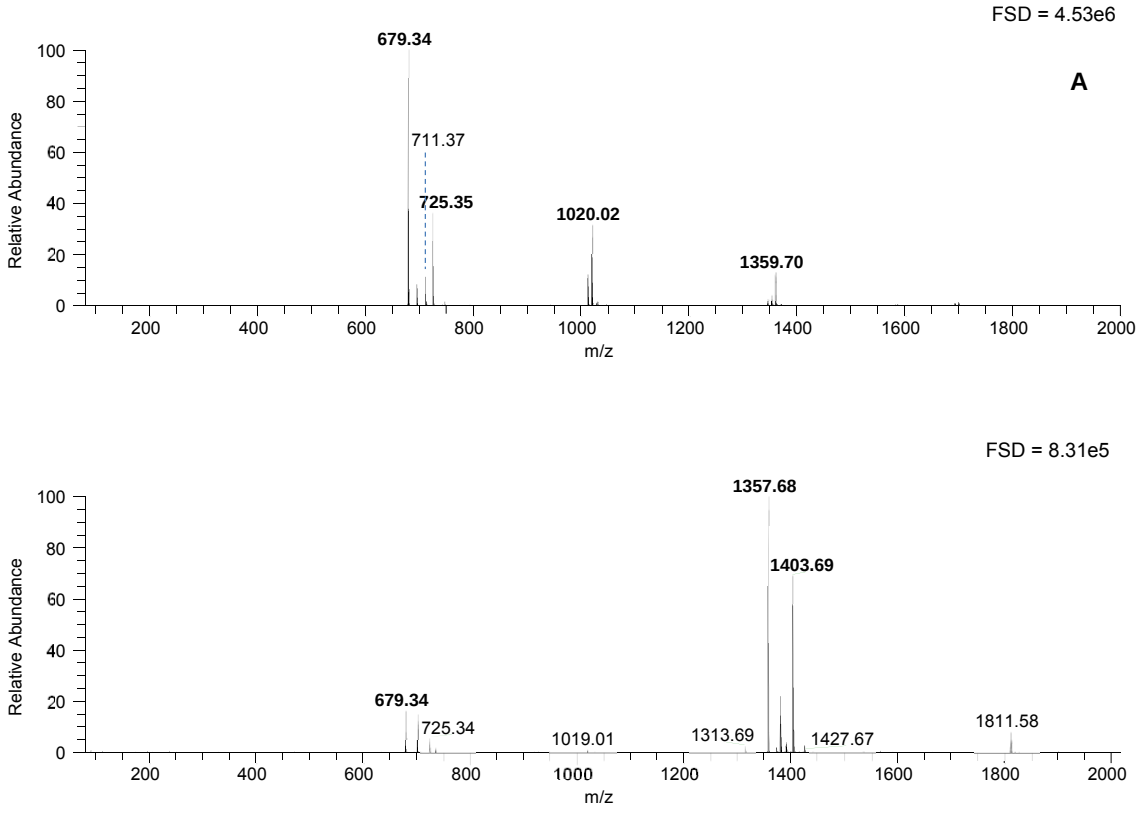


Figure 6. Anti-genotoxic effect of triterpenoids on H_2O_2-induced DNA damage. Anti-genotoxic effects of 100 nM CDDO and TRF after 24 hr pre-incubation on DNA damage in HT29 and CCD 841 cells challenged with 75 μM H_2O_2 and 25 μM H_2O_2 respectively. Data is presented as mean of 3 independent experiments ± SD compared to the untreated cells as control. One-way ANOVA and Post Hoc test Dunnett's T * $p < 0.05$.

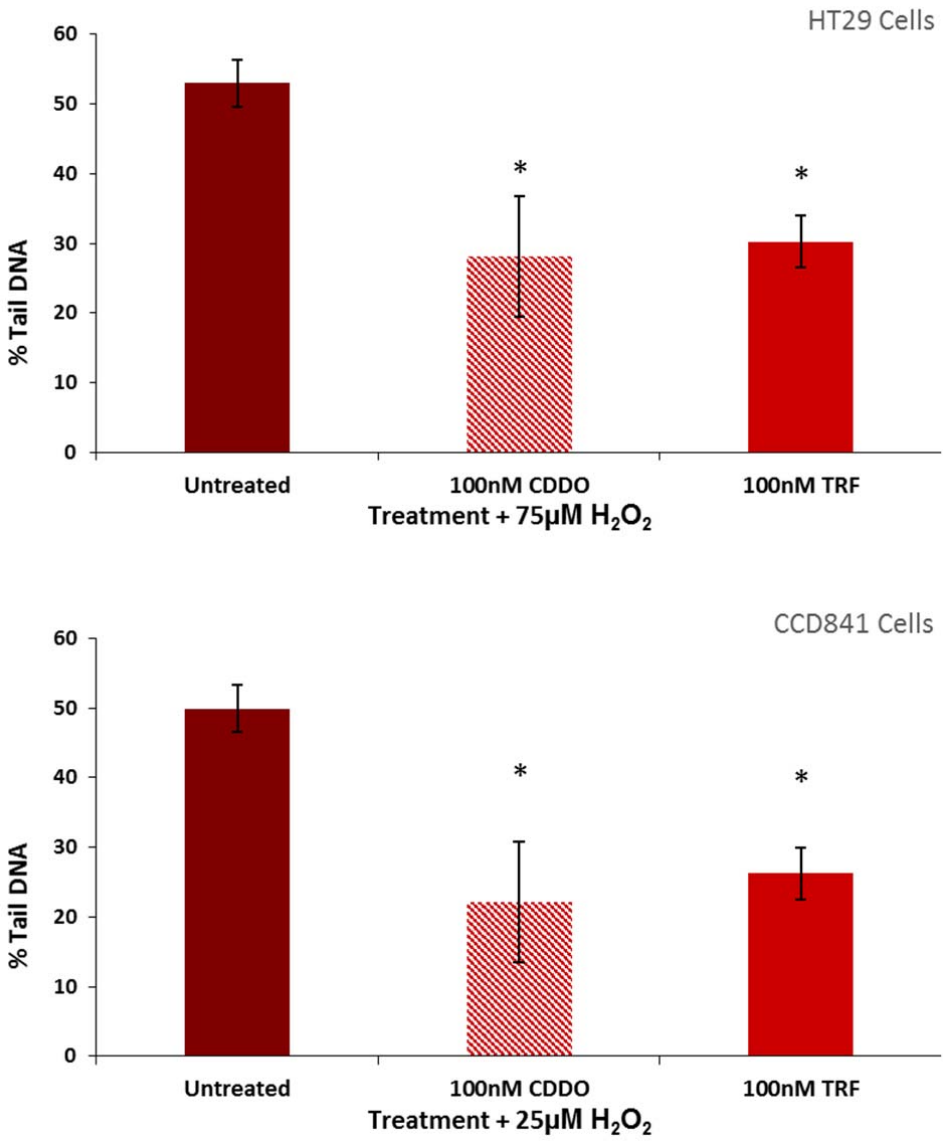


Figure 7. Effect of triterpenoids on antioxidant signalling pathway gene expression. Graphs show absolute fold change values (i.e. change in gene expression when compared to normalised untreated cell as control). Data is presented as the mean of 3 individual experiments (n=3) ± SD. Student t-test . * p<0.05,** p<0.01, *** p<0.001, **** p<0.001.

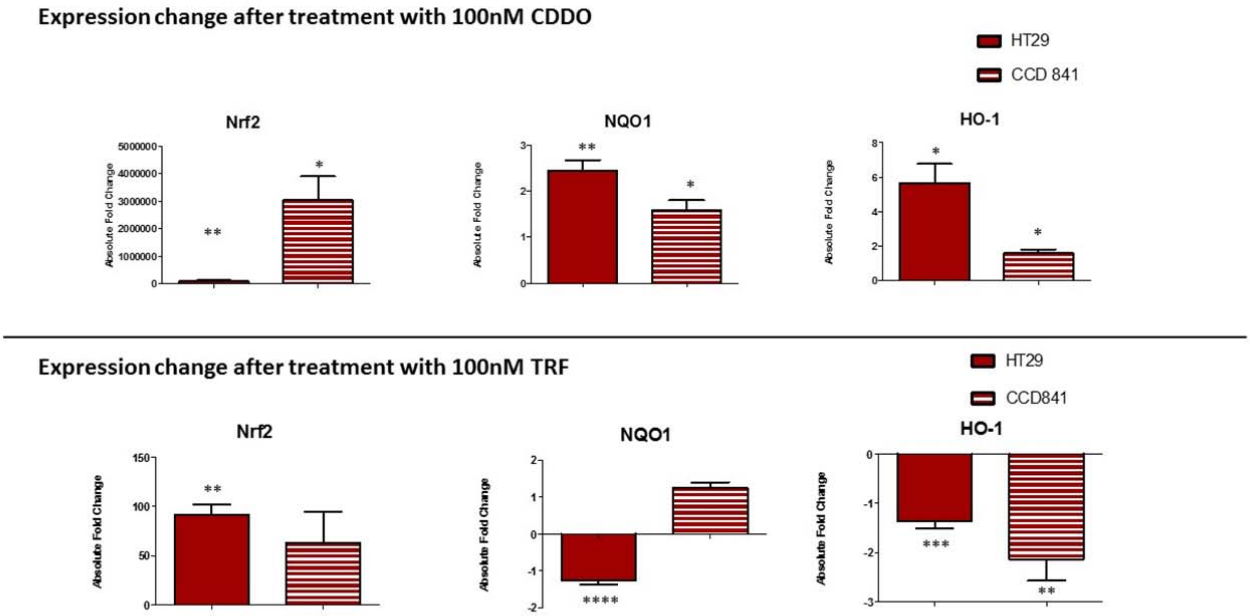


Table 1. Properties of selected unknown components in ileal samples

Peak	[M-H]$^-$ *(m/z)*	MS^2	Predicted formula	Putative identity
***m/z* 355** Negative	355.0648	337, ***209***, *191*	$C_{15}H_{15}O_{10}$	Caffeoyl glucuronide (http://www.hmdb.ca/metabolites/HMDB41705) but MS^2 data does not fit
  Positive	357.0809$^+$	339, ***147***	$C_{15}H_{17}O_{10}$ ($C_{15}H_{15}O_9$, $C_9H_7O_2$)[b]	*p*-coumaroyl glucarate [39, 40]
***m/z* 679** Negative	679.3648*	661, 559, 541, ***517***, 499, 455,	$C_{36}H_{55}O_{12}$	Components matching with $C_{36}H_{56}O_{12}$ formula were all triterpenoid glycoside derivatives based on oleanolic acid structures Tenuifolin (http://www.chemspider.com/Chemical-Structure.10205970.html?rid=52e5e4b2-c256-4cea-b530-813b9921295d)
  Positive	1359.7403	No MS^2	$C_{72}H_{111}O_{24}$	Trachelosperoside A1 (http://www.chemspider.com/Chemical-Structure.10273095.html) Suavissimoside F1 (http://www.chemspider.com/Chemical-Structure.10251895.html
	681.3833^{*+}	No MS^2	$C_{36}H_{57}O_{12}$	Esculentoside P (http://www.chemspider.com/Chemical-Structure.4956381.html?rid=661d5c11-37e3-4fdd-80c0-f9aaddbb2e34
	1361.7603$^+$	No MS^2	$C_{72}H_{113}O_{24}$	Ilexoside XLV [44] No components matched the $C_{72}H_{112}O_{24}$ formula

[a] Δ ppm < 2 for all formulae. MS^2 fragments in bold italics = most abundant signals. *****- $[M-H]^{2-}$ ion, has 0.5 amu variants, **+** = positive mode data, [b]formula in brackets are for the MS^2 fragments.

Table 2: Properties of components observed during purification

Peak	[M-H] (*m/z*)	MS^2	Predicted formula	Putative identity
T1	679.3451***** (725.3495)	661.3, 559.2, 541.4, ***517.3***, 499.3, 455.2, 437.3	$C_{36}H_{55}O_{12}$	Triterpenoid glycoside (+formate)
	1020.0237*****	1358.9, ***679.3***, 517.3, 455.2	None	Unknown
	1359.7463**	679.3	$C_{72}H_{111}O_{24}$	Triterpenoid glycoside dimer
T2	1357.6812** (1403.7355)	1313.2, 1151.2, 1018.7, ***679.3*****,** 517.1, 455.2	None	Triterpenoid glycoside dimer (+formate)
	679.3441*	541.3, ***517.4***, 499.3, 455.3	$C_{36}H_{55}O_{12}$	Triterpenoid glycoside
1	191.0130	***172.9***, 111.01	$C_6H_7O_7$	Citric acid
2	575.0983	557.1, 449.2, ***423.1*****,** ***407.0***, 289.2	$C_{30}H_{23}O_{12}$	A-type EC dimer
3	577.1143	559.3, 451.2, ***425.2***, 407.2, *289.2*	$C_{30}H_{25}O_{12}$	B-type EC dimer
4	121.0252	93.1	$C_7H_5O_2$	Benzoic acid
5	289.0612 (335.652)	***245.1***, 205.1, 179.1	$C_{15}H_{13}O_6$	EC (+formate)
6	561.1202 (607.1239)	543.1, 435.1, ***407.0***, 289.1	$C_{30}H_{25}O_{11}$	EfEC dimer (+formate)
7	849.1729	723.3, ***561.1***, 407.2, 289.1	$C_{45}H_{37}O_{17}$	Ef EC EC trimer
8	934.0376*	*1566.9*, *1234.9*, 897.0, 633.1, 301.3	None	Sanguiin H-6
9	833.1788	815.1, 707.2, ***561.0***, 289.2	$C_{45}H_{37}O_{16}$	EfEfEC trimer
10	833.1788	815.0, 707.2*,* ***561.0*****,** 289.1	$C_{45}H_{37}O_{16}$	EfEfEC trimer
11	1105.2369	979.1, 951.1, ***833.1***, 815.1, 707.3, *561.1*, 543.1	$C_{60}H_{49}O_{21}$	EfEfEfEC tetramer
12	147.0400	***118.9***, 84.9	$C_9H_7O_2$	Cinnamic acid

13	709.3553	663.2, ***501.3***	$C_{37}H_{57}O_{13}$	Formate adduct of triterpenoid glycoside [60]
14	709.3553	663.2, ***501.3***	$C_{37}H_{57}O_{13}$	Formate adduct of triterpenoid glycoside [60]
A	725.3490	***679.3***, 517.3	$C_{37}H_{57}O_{14}$	Formate adduct of triterpenoid glycoside
B	741.3428	***695.3***, 533.2	$C_{37}H_{57}O_{15}$	Formate adduct of unknown triterpenoid glycoside
C	493.2106	447.3, ***315.1***	$C_{22}H_{37}O_{12}$	Unknown
D	327.2056	309.3, 291.3, 229.2, ***171.1***	$C_{18}H_{31}O_5$	Trihydroxyoctadecadiene derivative (CS

All predicted formula derived with < 2 Δ ppm mass accuracy data; *$[M-H]^{2-}$ ion, has 0. 50 amu variants, **$[M-H]^{2-}$ and $[M-H]^{3-}$ions, has both 0.50 and 0.33 amu variants. MS^2 fragments in bold italics = most abundant signals

Shaded entries are the main triterpernoid peaks or are peaks enriched during purification, Underlined = major MS^2 fragments. Ef = epiafzelechin; EC = epicatechin.